Partial Kelly Portfolios and Shrinkage Estimators

Justin K. Rising
Department of Statistics
The Wharton School
University of Pennsylvania
Philadelphia, Pennsylvania
email: jrising@wharton.upenn.edu

Abraham J. Wyner
Department of Statistics
The Wharton School
University of Pennsylvania
Philadelphia, Pennsylvania
email: ajw@wharton.upenn.edu

***Abstract*—The log-optimal or Kelly portfolio forms the basis of a theoretically appealing investment strategy. However, it is difficult to compute, and this hinders its adoption in practice. In this paper we consider an approximate Kelly portfolio based on maximizing the expected value of a quadratic approximation to log utility. We show that this semi-log approximation gives an information-theoretic justification for portfolio selection based on either the mean-variance efficient frontier or the Sharpe ratio. We further show that there is a strong connection between estimated approximate fractional Kelly portfolios and shrinkage estimators, which leads to an optimal choice of a fractional Kelly parameter. We conclude by showing that the fractional Kelly portfolio succeeds not because of reduced risk, but because of reduced estimation error. We simulate to show that this property is largely responsible for the good empirical performance of fractional Kelly strategies.**

I. Introduction

The log-optimal or Kelly portfolio is a powerful tool but is not always easy to apply. On the one hand, under fairly general assumptions it is guaranteed to outperform any other portfolio in the long run. On the other hand, it is difficult to compute, unappetizingly volatile, and not guaranteed to outperform any specific portfolio over more modest time horizons. An alternative, due to [1], is the Universal Portfolio strategy, which will asymptotically converge to the log-optimal portfolio for an sequence of returns. However, it can be difficult to compute (although there are approximations) and the rate of convergence to the optimum is slow.

Another popular alternative is the fractional or partial Kelly strategy, which specifies that an investor puts a fraction α of their wealth into the Kelly portfolio, leaving the rest in cash while rebalancing in each investment period. While the fractional Kelly portfolio is no easier to compute than the full Kelly portfolio, it has one significant advantage: the volatility of the α-Kelly portfolio is simply α^2 times the volatility of the full Kelly portfolio, so even a modest partial Kelly strategy can significantly cut the investor's realized volatility. In theory, this decrease in volatility comes with a corresponding decrease in the long-run growth rate. In practice, however, partial Kelly strategies often outperform full Kelly strategies, and not just over very short horizons.

While this apparent contradiction is not of concern to the average investor, it does present a theoretical difficulty. In this paper, we offer a statistical resolution for the contradiction that applies in both the log-normal model for stock returns and in any model where the mean and variance of the market offer a good approximation of its performance. In breif, the operational procedure is simple: given a utility function and the estimated parameters of a model, it is possible, although sometimes computational hard, to estimate an optimal portfolio. However, the converse is true as well: if we are given a utility function and an estimated optimal portfolio, we can deduce the corresponding parameter estimates of for the model. We apply this observation to dissect an approximate fractional Kelly portfolio to discover that it is identical to the full Kelly portfolio constructed from an estimate of the mean return which is statistically shrunk towards the risk-free rate. Therefore, a fractional Kelly investor is in fact a full Kelly investor who uses an improved estimator of the market's parameters. There is power in this observation: while we do not know how to select an optimal Kelly fraction, we do know how to select an optimal shrinkage intensity. As these are the same problem, we can offer a principled and completely data-dependent method to select a Kelly fraction.

The remainder of this paper is laid out as follows. In Section II, we give a brief overview of shrinkage estimation and its applications in portfolio selection. In Section III, we lay out the basics of the market model and the expected utility framework in which we will work. In Section IV, we introduce the approximate Kelly portfolio, and show how this relates to the expected utility framework and the general framework of modern portfolio theory. In Section V, we exhibit a relationship between approximate fractional Kelly portfolios and shrinkage estimators, and argue that this relationship explains why fractional Kelly strategies work well in practice. Section VI briefly recapitulates the argument of the prior sections and lists some open problems.

II. A Brief Overview of Shrinkage Estimation

The theory of shrinkage estimation was introduced in [2], when Charles Stein observed that the sample average is an inadmissible estimator for the mean of a multivariate normal distribution in three or more dimensions. Since then it has become a major part of modern statistical theory and practice. We refer the interested reader to [3, 4] for an overview of the early approaches to and applications of shrinkage estimation.

The general recipe for a shrinkage estimator is simple. We are given two estimators: one with low variance and high bias, and one with low bias and high variance. We then form a

978-1-4673-2579-0/12/$31.00 ©2012 IEEE

convex combination of these two estimators which minimizes the expected squared ℓ^2 distance of the combination from the true value.

The most straightforward application of shrinkage estimation to portfolio selection is simply the use of improved estimators to reduce the effect of estimation error on the chosen portfolio. While some authors focus their attention solely on improving the sample mean (see [5]), the majority of effort is directed towards improving estimates of the sample covariance matrix as well (see, e.g., [6]–[8]).

However, there is evidence of a deeper connection between shrinkage estimation and portfolio selection. In [9], it is shown that picking the optimal long-only portfolio is equivalent to picking an unconstrained portfolio using a shrinkage estimator of the covariance matrix. In what follows, we will exhibit another such connection.

III. Preliminaries

A. A Simple Market Model

We begin with the strongest set of market assumptions that yield an interesting problem. We assume that the market consists of k risky assets whose return (the percentage change in price) in the nth period is denoted $\boldsymbol{R}_n$. When no period need be specified, we drop the subscript n. We further assume that there is one riskless asset whose constant return is denoted r_f. We will use $\boldsymbol{r_f}$ to denote the vector $r_f\vec{1}$.

We assume that the sequence of returns is independent and identically distributed with mean $\boldsymbol{\mu} \neq \boldsymbol{r_f}$ and invertible covariance matrix $\boldsymbol{\Sigma}$. We assume that any asset may be bought or sold in any quantity at any time without transaction costs, and that there are no restrictions on short sales.

We will use $\boldsymbol{w}$ to denote the weights of a portfolio comprised of the risky assets, and $G_i(\boldsymbol{w})$ to denote the equivalent continuously compounded interest rate during period n. A classic result of financial information theory (see [10]) states that

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} G_i(\boldsymbol{w}) = \left(1 - \boldsymbol{w}^T\vec{1}\right) r_f + \mathbb{E}\log(1+\boldsymbol{w}^T\boldsymbol{R})$$

which we will denote as $G(\boldsymbol{w})$. The Kelly portfolio $\boldsymbol{w}^*$ is simply the portfolio which maximizes $G(\boldsymbol{w})$. If we define our loss function to be

$$\ell(\boldsymbol{w}^*, \boldsymbol{w}) = \log(1+\boldsymbol{w}^{*T}\boldsymbol{R}) - \log(1+\boldsymbol{w}^T\boldsymbol{R})$$

our risk function is then given by

$$L(\boldsymbol{w}^*, \boldsymbol{w}) = G(\boldsymbol{w}^*) - G(\boldsymbol{w})$$

This is simply the amount by which the growth rate of the optimal portfolio exceeds that of the portfolio with weights $\boldsymbol{w}$.

Finally, we will write $X_{\boldsymbol{w}}$ to denote the total return to the portfolio $\boldsymbol{w}$ in a single period. By definition, we have

$$X_{\boldsymbol{w}} = (1-\boldsymbol{w}^T\vec{1})r_f + \boldsymbol{w}^T\boldsymbol{R}$$

but our future calculations will be easier to interpret if we use the equivalent formulation

$$X_{\boldsymbol{w}} = r_f + \boldsymbol{w}^T(\boldsymbol{R} - \boldsymbol{r_f})$$

B. Risk Aversion and Utility

An investor's tolerance for risk is quantified with a utility function. Not all choices are are unambiguous : any two utility functions U_1 and U_2 such that $U_2(x) = aU_1(x)+b$ with $a>0$ will lead to the same relative decisions in all circumstances (see, e.g., [11]). Therefore, we wish to have a measure of risk tolerance which is invariant under affine transformations. There is a ready-made solution in the Arrow-Pratt risk aversion measures defined in [12, 13]. We will focus on the relative risk aversion, which is defined by

$$r(x) = -\frac{xU''(x)}{U'(x)}$$

Much can be written about the interpretation of this quantity, but we will be satisfied with the observation that if we imagine two investors with the same total wealth, the one with the higher relative risk aversion will prefer to keep a larger proportion of their wealth in a safe asset.

Under the assumption that the returns in each period are independent and identically distributed, an investor who invests in the Kelly portfolio $\boldsymbol{w}^*$ will invest a constant proportion of wealth in each period. As one might expect, this is optimal in the expected utility framework exactly when the investor's relative risk aversion is constant (see [14, 15]). As shown in [12, 13], the only utility functions with constant relative risk aversion are the isoelastic utility functions, which are of the form

$$U_\theta(X_{\boldsymbol{w}}) = \begin{cases} \dfrac{(1+X_{\boldsymbol{w}})^{1-\theta}-1}{1-\theta} & \theta \neq 1 \\ \log(1+X_{\boldsymbol{w}}) & \theta = 1 \end{cases}$$

An investor whose utility function is equal to U_θ will have relative risk aversion identically equal to θ. We briefly note that $\theta \leq 0$ does not correspond empirically to investor behavior (see [13, 16]). As a result, we assume through the remainder of this paper that $\theta > 0$.

IV. Approximate Fractional Kelly Portfolios and Modern Portfolio Theory

In this section, we assume that the distribution of returns is known. We will give a simple formula for an approximation to the fractional Kelly portfolio, and reinterpret standard concepts in portfolio theory in light of our findings.

Following [17]–[20], we apply Taylor's theorem to find that

$$U_\theta(X_{\boldsymbol{w}}) = X_{\boldsymbol{w}} - \frac{\theta}{2}X_{\boldsymbol{w}}^2 + O(X_{\boldsymbol{w}}^3)$$

and

$$\mathbb{E}U_\theta(X_{\boldsymbol{w}}) = r_f + w^T(\boldsymbol{\mu} - \boldsymbol{r_f}) - \frac{\theta}{2}\mathbb{E}(X_{\boldsymbol{w}})^2 + O(\mathbb{E}(X_{\boldsymbol{w}}^3))$$

We assume that

$$\mathbb{E}(X_{\boldsymbol{w}})^2 \approx \mathrm{Var}\,(X_{\boldsymbol{w}})$$

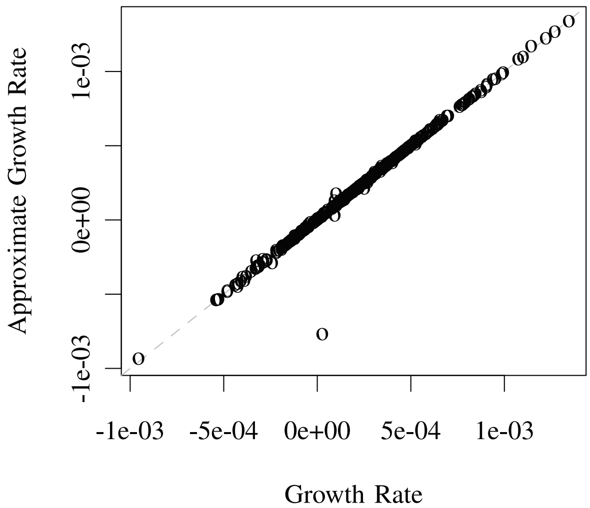

Fig. 1. Expected log utility versus the quadratic approximation for each component stock of the S&P 500 which has traded continuously under the same ticker symbol from January 1st, 1999 to December 31st, 2010. The single outlier is Pier One (PIR), whose stock price nearly quadrupled on March 23, 2009. The maximum absolute deviation is approximately 8×10^{-4}, and the mean absolute deviation excluding PIR is approximately 3.9×10^{-6}.

and

$$\mathbb{E}(X_{\boldsymbol{w}})^n \approx 0$$

for all $n \geq 3$, so we define

$$V_\theta(\boldsymbol{w}) = r_f + \boldsymbol{w}^T(\boldsymbol{\mu} - \boldsymbol{r_f}) - \frac{\theta}{2}\boldsymbol{w}^T\boldsymbol{\Sigma}\boldsymbol{w}$$

as our approximation to expected utility. We then have that $G(\boldsymbol{w}) \approx V_1(\boldsymbol{w})$. As shown in Figure 1, the approximation is empirically quite accurate for large cap stocks.

The Kelly portfolio weights $\boldsymbol{w}^*$ are exactly equal to $\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - \boldsymbol{r_f})$ when returns are normally distributed ([21]). A simple derivative argument shows that this is approximately true more generally:

Theorem IV.1. *The approximate Kelly portfolio weights* $\tilde{\boldsymbol{w}}^*$ *are equal to* $\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - \boldsymbol{r_f})$.

Proof: We have that $\nabla_{\boldsymbol{w}} V_1(\boldsymbol{w}) = \boldsymbol{\mu} - \boldsymbol{r_f} - \boldsymbol{\Sigma} w$, and that $\nabla_{\boldsymbol{w}} V_1(\tilde{\boldsymbol{w}}^*) = 0$. This implies immediately that $\tilde{\boldsymbol{w}}^* = \boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - \boldsymbol{r_f})$. $\nabla^2_{\boldsymbol{w}} V_1(w) = -\boldsymbol{\Sigma}$, which is always negative definite, and so we have that $\tilde{\boldsymbol{w}}^*$ maximizes $V_1(\boldsymbol{w})$. ∎

We can carry the proof of Theorem IV.1 through for general θ to derive the following corollary:

Corollary IV.2. *The* $\frac{1}{\theta}$*-Kelly portfolio approximately optimizes* $\mathbb{E}U_\theta(X_{\boldsymbol{w}})$.

Corollary IV.2 is qualitatively the same result as is found in [22] for Kelly bets.

We pause to comment on two standard topics in portfolio theory. The first is portfolio selection based on the mean-variance efficient frontier detailed in [23, 24]. In brief, the efficient frontier consists of the set of portfolios w which maximize functions of the form $\boldsymbol{w}^T(\boldsymbol{\mu} - \boldsymbol{r_f}) - \frac{\theta}{2}\boldsymbol{w}^T\boldsymbol{\Sigma}\boldsymbol{w}$ where $\theta \geq 0$ is a free parameter measuring risk tolerance. These functions are exactly our approximations to expected isoelastic utility, and so we have an alternate characterization of the efficient frontier. While the mean-variance approximation to logarithmic utility is well-known (see [17]–[20]), the use of the efficient frontier to approximate the expectations of all isoelastic utilities is not. Furthermore, we can interpret the risk tolerance parameter θ as the Arrow-Pratt relative risk aversion discussed in Section III-B.

The second is the Sharpe ratio s ([25, 26]). The maximum possible portfolio growth rate is equal to

$$\frac{1}{2}(\boldsymbol{\mu} - \boldsymbol{r_f})^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - \boldsymbol{r_f})$$

which is the squared Mahalanobis distance between the mean return and the vector $\boldsymbol{r_f}$. In the case where there is a single risky asset (which may itself be a portfolio of assets), this is simply $\frac{1}{2}s^2$. Therefore, a portfolio formed from an asset with a high squared Sharpe ratio will eventually outperform a portfolio formed from an asset with a low squared Sharpe ratio, and an investor who is selecting a single asset for investment is justified in using the Sharpe ratio on information-theoretic grounds. Furthermore, an investor who is selecting a collection of assets should examine the maximum growth rate, not the individual Sharpe ratios, as they do not account for the correlations between assets.

We close this section with a brief discussion of the relationship between growth, volatility and fractional Kelly portfolios. These portfolios are known to offer an effective tradeoff between long term growth and short term security according to several measures ([27, 28]), but they are not known to offer an optimal tradeoff. We now show that approximate fractional Kelly portfolios do minimize a weighted combination of approximate lost growth and return variance:

Theorem IV.3. *The* $\frac{1}{2\lambda+1}$*-partial Kelly portfolio minimizes the quantity* $V_1(\tilde{\boldsymbol{w}}^*) - V_1(\boldsymbol{w}) + \lambda \operatorname{Var}(X_{\boldsymbol{w}})$.

Proof: We have that

$$\nabla_{\boldsymbol{w}}\left(V_1(\tilde{\boldsymbol{w}}^*) - V_1(\boldsymbol{w}) + \lambda \operatorname{Var}(X_{\boldsymbol{w}})\right) = \boldsymbol{\mu} - \boldsymbol{r_f} + (2\lambda - 1)\boldsymbol{\Sigma}\boldsymbol{w}$$

This is equal to zero exactly when $\boldsymbol{w} = \frac{1}{2\lambda+1}\boldsymbol{w}^*$ as claimed. The Hessian is positive definite for all $\boldsymbol{w}$ as long as $\lambda > -\frac{1}{2}$, so this point is a global minimum. ∎

V. Approximate Fractional Kelly Portfolios and Shrinkage Estimators

We now consider the more realistic case where the distribution of returns is unknown. We will exhibit a relationship between estimated approximate fractional Kelly portfolios and shrinkage estimators, and argue by simulation that this connection explains the empirical success of fractional Kelly portfolios.

Let $\hat{\boldsymbol{\mu}}$ denote the sample mean, and let $\hat{\boldsymbol{\Sigma}}^{-1}$ denote the sample precision matrix. We proceed as before, with all distribution-dependent quantities replaced by their plug-in estimates. The approximate growth rate is estimated by

$$\hat{V}_1(\boldsymbol{w}) = \boldsymbol{w}^T\hat{\boldsymbol{\mu}} - \frac{1}{2}\boldsymbol{w}^T\hat{\boldsymbol{\Sigma}}\boldsymbol{w}$$

and the approximate Kelly portfolio is estimated by

$$\hat{\tilde{\boldsymbol{w}}}^* = \hat{\boldsymbol{\Sigma}}^{-1}\left(\hat{\boldsymbol{\mu}} - \boldsymbol{r_f}\right)$$

While estimation error is detrimental to portfolio selection in general, there is strong evidence that errors in estimating the mean return are much worse than errors in estimating the variances or covariances (see [29] for a survey of the relevant literature), and that overestimating the excess mean return is significantly worse than underestimating the same quantity (see Section 7.3 of [21]). We should therefore expect a portfolio formed from an improved estimator of the mean to be superior to a portfolio formed from the sample mean. The standard method of improving an estimator is to shrink it toward some target, but the selection of an appropriate target is an art in and of itself. Fortunately, we have been given an easy prospective target: because the magnitude of the approximate Kelly portfolio weights is driven by the difference $\boldsymbol{\mu} - \boldsymbol{r_f}$, we can shrink $\hat{\boldsymbol{\mu}}$ towards $\boldsymbol{r_f}$ to avoid over investment in any particular asset.

We will form our portfolio from an estimator of the form

$$\alpha\boldsymbol{r_f} + (1-\alpha)\hat{\boldsymbol{\mu}}$$

The approximate Kelly portfolio associated with this estimator is given by

$$(1-\alpha)\hat{\boldsymbol{\Sigma}}^{-1}\left(\hat{\boldsymbol{\mu}} - \boldsymbol{r_f}\right)$$

We see immediately that we could have selected the same portfolio as an approximate fractional Kelly portfolio based on the sample mean. We capture this observation in Theorem V.1:

Theorem V.1. *The full Kelly portfolio formed from an estimator of the form $\alpha\boldsymbol{r_f} + (1-\alpha)\hat{\boldsymbol{\mu}}$ with $0 \le \alpha \le 1$ is identical to the $(1-\alpha)$-Kelly portfolio formed from $\hat{\boldsymbol{\mu}}$.*

While it is not generally clear how we should choose among fractional Kelly portfolios, we can always form shrinkage estimators by minimizing the expected squared distance from the true parameter. Our next theorem gives an optimal shrinkage procedure that we can use in portfolio selection:

Theorem V.2. *Let $\hat{\boldsymbol{\mu}}$ denote the sample mean, and let*

$$\alpha_* = \operatorname*{argmin}_{\alpha\in[0,1]} \mathbb{E}||\boldsymbol{\mu} - \alpha\boldsymbol{r_f} - (1-\alpha)\hat{\boldsymbol{\mu}}||$$

Then under the assumption that the distribution of returns has finite second moment, we have that

$$\alpha_* = \frac{\operatorname{tr}\left(\boldsymbol{\Sigma}\right)}{n||\boldsymbol{\mu} - \boldsymbol{r_f}||_2^2} + O\left(\frac{1}{n^2}\right)$$

and we may estimate α_* by

$$\hat{\alpha}_* = \frac{\operatorname{tr}\left(\hat{\boldsymbol{\Sigma}}\right)}{n||\hat{\boldsymbol{\mu}} - \boldsymbol{r_f}||_2^2}$$

Proof: Our proof is adapted from a more general result in [7], we provide a sketch. We are given an expression for squared loss that depends upon a parameter α, and so we perform standard derivative calculations to find that

$$\alpha_* = \frac{\operatorname{tr}\left(\operatorname{Var}\left(\hat{\boldsymbol{\mu}}\right)\right)}{\operatorname{tr}\left(\operatorname{Var}\left(\hat{\boldsymbol{\mu}}\right)\right) + ||\boldsymbol{\mu} - \boldsymbol{r_f}||_2^2}$$

Standard asymptotics then allow us to conclude that

$$\alpha_* = \frac{1}{n}\cdot\frac{\operatorname{tr}\left(\operatorname{AsyVar}\left(\hat{\boldsymbol{\mu}}\right)\right)}{||\boldsymbol{\mu} - \boldsymbol{r_f}||_2^2} + O\left(\frac{1}{n^2}\right)$$

The asymptotic variance of $\hat{\boldsymbol{\mu}}$ is simply $\boldsymbol{\Sigma}$, so we have that

$$\alpha_* = \frac{1}{n}\cdot\frac{\operatorname{tr}\left(\boldsymbol{\Sigma}\right)}{||\boldsymbol{\mu} - \boldsymbol{r_f}||_2^2} + O\left(\frac{1}{n^2}\right)$$

as claimed. $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$ are consistent estimators of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ respectively, so if we plug them in to our formula for α_*, we receive a consistent estimator $\hat{\alpha}_*$. ∎

The theory we have laid out in this section gives us reason to believe that the empirical success of the fractional Kelly strategy is due to its approximation of an improved mean estimator. We close with a simulation that offers evidence for this claim. We take the same set of returns as those used in Figure 1, and consider portfolios formed from twenty-five randomly selected stocks. We compute the growth lost due to estimation error for two portfolios:

1) The portfolio formed from the improved mean estimator $\hat{\alpha}_*\boldsymbol{r_f} + (1-\hat{\alpha}_*)\hat{\boldsymbol{\mu}}$ and the sample covariance matrix $\hat{\boldsymbol{\Sigma}}$.
2) The portfolio formed from the sample mean $\hat{\boldsymbol{\mu}}$ and the true covariance matrix $\boldsymbol{\Sigma}$.

Figure 2 shows the risks associated with these two portfolios across 1000 trials. It is clear that the first portfolio far outperforms the second, even though the second is computed from the true covariance matrix. In light of these results, we assert that improved mean estimation is largely responsible for the good empirical performance of partial Kelly portfolios.

VI. Conclusion and Open Problems

In Theorem V.1, we have shown that the partial Kelly portfolio formed from the sample mean is identical to the full Kelly portfolio formed from a shrinkage estimator of the population mean. In Theorem V.2 we have shown how to optimally pick the proportion of wealth to be invested in the Kelly strategy. While these results are noteworthy in and of itself, there is an interesting idea here: every combination of utility function and optimal portfolio implies parameters for the market.

We have also shown a connection between Kelly portfolio theory and several key concepts in modern portfolio theory. Our approximate Kelly portfolios trace out the mean-variance efficient frontier, and grow at a rate which is closely tied to a generalization of the Sharpe ratio. This last part is compelling, as the Sharpe ratio is an intuitive measure of portfolio performance, but has not been placed on a firm basis.

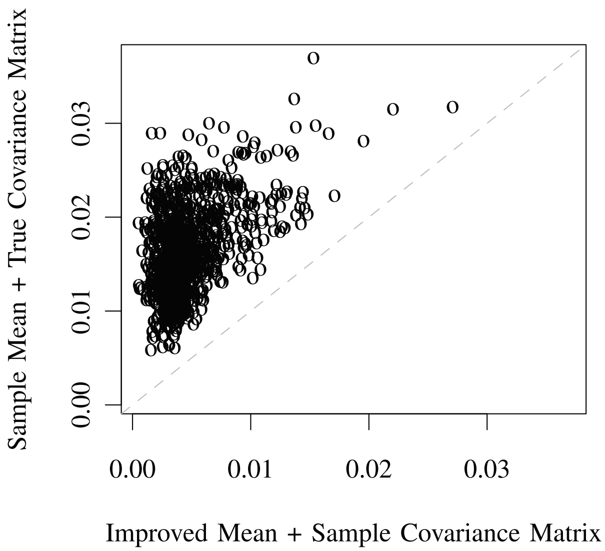

Fig. 2. Approximate growth lost from two portfolios with estimation error on simulated returns series.

We close with two open problems related to Kelly strategies and shrinkage estimation. The first relates to the role of covariance matrix shrinkage in portfolio selection. While estimation error in the mean is the most important quantity to control, it is clear that estimation error in the covariance matrix is also detrimental, and as mentioned earlier, a large body of literature exists which uses improved covariance matrix estimates to construct improved portfolios. In future work we will extend our theory to give insight as to how the estimated covariance matrix should be shrunk.

The second corresponds to Kelly gambling. As in the portfolio selection problem, we see an analogy between partial Kelly strategies and shrinkage estimation of the gambler's edge. However, it is more difficult to make this mathematically precise because the third and higher moments of the returns are not negligible as they are in daily stock prices.

References

[1] T. M. Cover, "Universal Portfolios," *Mathematical Finance*, vol. 1, no. 1, pp. 1–29, 1991.

[2] C. Stein, "Inadmissibility of the Usual Estimator for the Mean of a Multivariate Normal Distribution," in *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, 1956, pp. 197–206.

[3] B. Efron and C. Morris, "Data Analysis Using Stein's Estimator and its Generalizations," *Journal of the American Statistical Association*, vol. 70, no. 350, pp. 311–319, 1977.

[4] ——, "Stein's Paradox in Statistics," *Scientific American*, vol. 236, no. 5, pp. 119–127, 1977.

[5] P. Jorion, "Bayes-Stein Estimation for Portfolio Analysis," *Journal of Financial and Quantitative Analysis*, vol. 21, no. 3, pp. 279–292, 1986.

[6] L. C. MacLean, M. Foster, and W. T. Ziembra, "Covariance Complexity and Rates of Returns on Assets," *Journal of Banking & Finance*, vol. 31, no. 11, pp. 3503–3523, 2007.

[7] O. Ledoit and M. Wolf, "Improved Estimation of the Covariance Matrix of Stock Returns With an Application to Portfolio Selection," *Journal of Empirical Finance*, vol. 10, no. 5, pp. 603–621, 2003.

[8] ——, "Honey, I Shrunk the Sample Covariance Matrix," *Journal of Portfolio Management*, vol. 30, no. 4, pp. 110–119, 2004.

[9] R. Jagannathan and T. Ma, "Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps," *The Journal of Finance*, vol. 58, no. 4, pp. 1651–1683, 2003.

[10] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Wiley-Interscience, 2006.

[11] R. B. Myerson, *Game Theory: Analysis of Conflict*. Harvard University Press, 1997.

[12] K. J. Arrow, "The Theory of Risk Aversion," in *Essays in the Theory of Risk Bearing*. Markham Publishing Company, 1971, pp. 90–109.

[13] J. W. Pratt, "Risk Aversion in the Small and in the Large," *Econometrica*, vol. 32, no. 1-2, pp. 122–136, 1964.

[14] P. A. Samuelson, "Lifetime Portfolio Selection by Dynamic Stochastic Programming," *Review of Economics and Statistics*, vol. 51, no. 3, pp. 239–246, 1969.

[15] J. Mossin, "Optimal Multiperiod Portfolio Policies," *Journal of Business*, vol. 41, no. 2, pp. 215–229, 1968.

[16] I. Friend and M. Blume, "The Demand for Risky Assets," *American Economic Review*, vol. 65, no. 5, pp. 900–922, 1975.

[17] C. Y. Belentepe and A. J. Wyner, "A Statistical View of Universal Stock Market Portfolios," in *IEEE International Symposium on Information Theory*, 2005, pp. 573–577.

[18] L. Györfi, A. Urbán, and I. Vajda, "Kernel-Based Semi-Log Optimal Empirical Portfolio Selection Strategies," *International Journal of Theoretical and Applied Finance*, vol. 10, no. 3, pp. 505–516, 2007.

[19] G. Ottucsák and I. Vajda, "An Asymptotic Analysis of the Mean-Variance Portfolio Selection," *Statistics & Decisions*, vol. 25, pp. 63–88, 2007.

[20] L. B. Pulley, "Mean-Variance Approximations to Expected Logarithmic Utility," *Operations Research*, vol. 31, no. 4, pp. 685–696, 1983.

[21] E. O. Thorp, "The Kelly Criterion in Blackjack, Sports Betting, and the Stock Market," in *The Kelly Capital Growth Investment Criterion: Theory and Practice*. World Scientific Publishing Company, 2011, pp. 789–832.

[22] J. B. Kadane, "Partial-Kelly Strategies and Expected Utility: Small Edge Asymptotics," *Decision Analysis*, vol. 8, no. 1, pp. 4–9, 2011.

[23] H. M. Markowitz, "Portfolio Selection," *Journal of Finance*, vol. 7, no. 1, pp. 77–91, 1952.

[24] H. Markowitz, *Portfolio Selection: Efficient Diversification of Investments*. John Wiley & Sons, 1959.

[25] W. F. Sharpe, "Mutual Fund Performance," *The Journal of Business*, vol. 39, no. 1, pp. 119–138, 1966.

[26] ——, "The Sharpe Ratio," *The Journal of Portfolio Management*, vol. 21, no. 1, pp. 49–58, 1994.

[27] L. C. MacLean, W. T. Ziemba, and G. Blazenko, "Growth versus Security in Dynamic Investment Analysis," *Management Science*, vol. 38, no. 11, pp. 1562–1585, 1992.

[28] L. C. MacLean and W. T. Ziemba, "Growth versus Security Tradeoffs in Dynamic Investment Analysis," *Annals of Operations Research*, vol. 85, pp. 193–225, 1999.

[29] L. C. MacLean, E. O. Thorp, and W. T. Ziemba, "Long-Term Capital Growth: the Good and Bad Properties of the Kelly and Fractional Kelly Capital Growth Criteria," in *The Kelly Capital Growth Investment Criterion: Theory and Practice*. World Scientific Publishing Company, 2011, pp. 563–574.